MDPI

Article

TEC forecasting based on manifold trajectories

Enrique Monte Moreno [1,*], Alberto García Rigo [2,3], Manuel Hernández-Pajares [2,3] and Heng Yang [1]

[1] Department of Signal Theory and Communications, TALP research center, Technical University of Catalonia, Barcelona, Spain
[2] UPC-IonSAT, Technical University of Catalonia, Barcelona, Spain
[3] IEEC-CTE-CRAE, Institut d'Estudis Espacials de Catalunya, Barcelona, Spain
* Correspondence: enric.monte@upc.edu; Tel.: +34-934016435

Academic Editor: name
Version June 8, 2018 submitted to Remote Sens.

Abstract: In this paper, we present a method for forecasting the ionospheric Total Electron Content (TEC) distribution from the International GNSS Service's Global Ionospheric Maps. The forecasting system gives an estimation of the value of the TEC distribution based on linear combination of previous TEC maps (i.e. a set of 2D arrays indexed by time), and the computation of a tangent subspace in a manifold associated to each map. The use of the tangent space to each map is justified because it allows to model the possible distortions from one observation to the next as a trajectory on the tangent manifold of the map. The coefficients of the linear combination of the last observations along with the tangent space are estimated at each time stamp in order to minimize the mean square forecasting error with a regularization term. The the estimation is made at each time stamp to adapt the forecast to short-term variations in solar activity..

Keywords: Total Electron Content; Ionosphere; Forecasting; Tangent Distance; GNSS

1. Introduction

The capability to monitor and forecast Space Weather in the near-Earth environment, in particular ionosphere, is becoming more and more critical nowadays. There are a number of applications that can benefit from the estimates of the nowcast and forecast ionospheric state. In particular, those relying on Global Navigation Satellite Systems (GNSS) and Satellite Based Augmentation Systems (SBAS).

In fact, a perturbed ionospheric state can affect multiple systems that rely on signals that propagate through ionospheric refractive media. Indeed, severe ionospheric disturbances can imply loss of lock in GNSS receivers located on Earth's surface or on-board LEO satellites, EGNOS performance degradation events [21], [2], degradation of precise positioning, among others. In addition, monitoring the near-Earth environment can also help on taking appropriate decisions by space launch operators ([4], [20]). In this context, one decade ago the European Space Agency (ESA) already showed their interest for two-day ahead predictions for mission planning, in particular, for the Soil Moisture and Ocean Salinity mission [3], [5], [1].

Several methods for the forecasting of ionospheric parameters estimated from GNSS data have been developed in the last years. For instance [6], the approach by ESA, based on the extrapolation of Spherical Harmonics coefficients, and the UPC approach by [1], based on the Discrete Cosine Transform (DCT). One method of estimating the Vertical Total Electron Content (VTEC) maps is to combine several predicted VTEC Global Ionospheric Maps (GIM), as it was done in the context of International GNSS Service (IGS) Ionosphere Working Group (IGS Iono-WG).

In addition, methods based on neural networks have also been developed [7], [8], [9], [23], autocorrelation and autocovariance procedures [10], [11], [12], linear regression [13], [14], the Grey

model by [22]. On the other hand, methods have been developed that depend on physical models, such as [24], which is the one considered by the JPL Global Ionosphere Thermosphere Model (GITM).

The TEC forecasting method relies on time series from the UPC Vertical TEC (VTEC) GIMs (labelled UQRG) computed by TOMION v1 software in the frame of the International GNSS Service (IGS; [36] and [15]). TOMION is a software mainly developed by one of the authors of this manuscript (MHP) consisting on two main versions: TOMION v1, which is focused on the estimation of electron content models, mostly based on GNSS dual-frequency measurements with ionospheric tomographic and kriging interpolation; and TOMION v2, which consists on a hybrid geodetic and tomographic ionospheric model that allows for precise ionospheric determination and precise positioning, among GPS radio-occultation data processing, see [15] , [18], [17], [16]. In particular, either UQRG 15-min resolution with 1-day latency or the UPC real-time VTEC maps (labelled URTG), which are computed continuously at a 30-seconds rate, can be ingested to the forecasting approach. In both cases, the GIMs provide bidimensional VTEC values assuming a thin single-layer ionospheric model at a height of 450 km and considering a world-wide regular grid every $5\,^{\circ}/2.5^{\circ}$ in longitude/latitude range (i.e. 71 latitudinal and 72 longitudinal points accounting for a total of 5112 VTEC values per map). Note that for building the forecasting time series, these maps are transformed to a local-time sun-fixed reference frame.

In brief, the way to calculate VTEC values in TOMION relies on the simultaneous estimation of the electron density values of the ionospheric electron content distribution, considering a basis functions of a partition in voxels, and the ambiguity of the ionospheric combination of carrier phases. And both set of parameters are estimated in a forward Kalman filter. Then, each Slant Total Electron Content (STEC) affecting each GNSS receiver-satellite signals can be estimated as $L_I = L_1 - L_2 \approx \alpha_I STEC + B_I$.

Where $\alpha_I = \alpha_2 - \alpha_1 \approx 1.05$ and $\alpha_i = 40.3/f_i^2$. B_I is the carrier-phase ionospheric ambiguity, which is estimated as part of the TOMION processing chain to derive UPC GIMs. In addition, Kriging interpolation technique [16] is used to fill the gaps where there is lack of data, given the inhomogeneous sparsity of GNSS receivers [25], [26].

The results presented in this work have been obtained from UQRG maps, since these have proven to have the best performance against independent TEC measurements - JASON test - from altimeter data, as well as independent GNSS receivers -dSTEC test- among the different IGS models (for the details, see [19]). In particular, the results from this paper have been derived from UQRG data in years 2014 -coinciding with the mild 24^{th} Solar Cycle maximum-, 2015 and 2016.

2. Justification of the method

The method that we propose is an evolution of a forecasting method based on a bi-dimensional Discrete Cosine Transform (2D-DCT) by [1] and [20]. Originally it was decided to use the DCT method by analogy with the video compression methods based on predictive coding (for instance, see chapter 8 of [27]).

The analogy lies in the fact that the time evolution of the TEC maps can be understood as a series of images, where each map is indexed by its temporal index. The suitability of the use of a 2D-DCT for describing the TEC maps is due to the smoothness of the maps, i.e. its low spacial frequency components. In other words, typically most of the energy ($> 95\%$) is concentrated in normalized horizontal and vertical frequencies less than $1/10$. This property allowed for the forecasting based on a small subset of coefficients on the DCT transform domain, from which a given map can reconstructed with low distortion. Therefore, instead of dealing with the whole maps, the method only modeled the time evolution of the subset of DCT coefficients that accounted for most of the temporal changes of the general shape of the ionized zones. That meant that the forecast was done using a basis extracted from the low frequency components of the DCT that spans the space of possible TEC global maps (in the linear algebra sense). The result is that the prediction is correct throughout the map except at the areas corresponding to transitions between different levels of ionization, that is at the borders of these high-ionization regions (mainly in the case of the Appleton-Hartree anomaly). Such regions occupied a

small area of the whole map, and specifically were the regions that contributed the most to the forecast error. Therefore, the temporal distortions of the shape of the areas with the highest ionization were not well modeled locally. The new forecasting model is inspired by this fact. That is, although these border regions represent a small fraction of the whole area/map being forecasted, they account for most of the error.

As an example, we present the forecasts at two different times of the year with different solar activity. We decided to select different dates in order to show that the error at the edges of the high ionization zones occurs in different levels of solar activity. Also to show, that the effect occurs to a lesser extent in the model we propose in this article. Figures 1 and 2 correspond to the prediction with the DCT method and Figure 8 to the tangent plane method presented in this article. Figure 1 corresponds to the month of May (lower solar activity, see Figure 13), the 2 to the month of February. Figure 8 was computed at the same dates.

The improvement due to the use of tangent planes is evident in the fact that the areas with an error of more than 20 TECUs (Total Electron Content, which is defined as 1 TECU=10^{16} el/m2) are located in much smaller areas of the error maps.

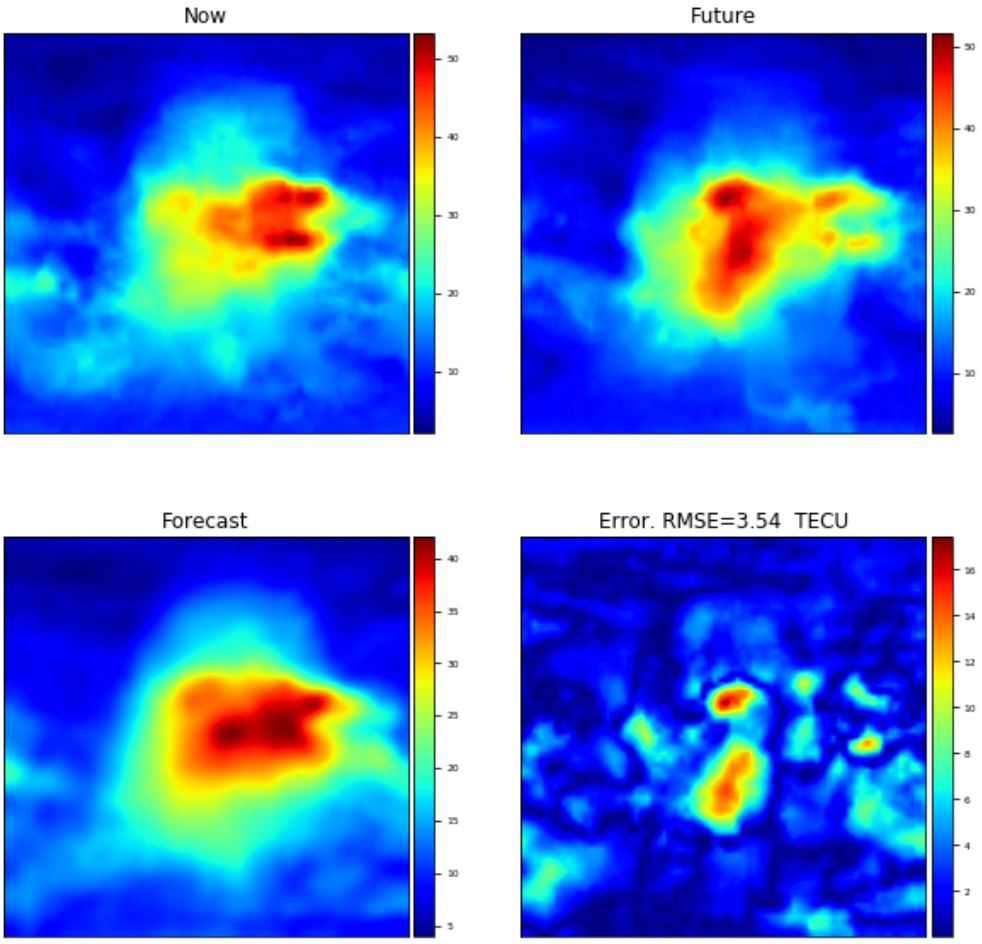

Figure 1. Forecasting by means of the DCT method at a horizon of 3 hours. Time stamp, 2016-05-12 15:15:00 (Year-Month-Day Hour:Minute:Second). Current map (upper left), future map (upper right), forecast (lower left), and forecast error (lower right).

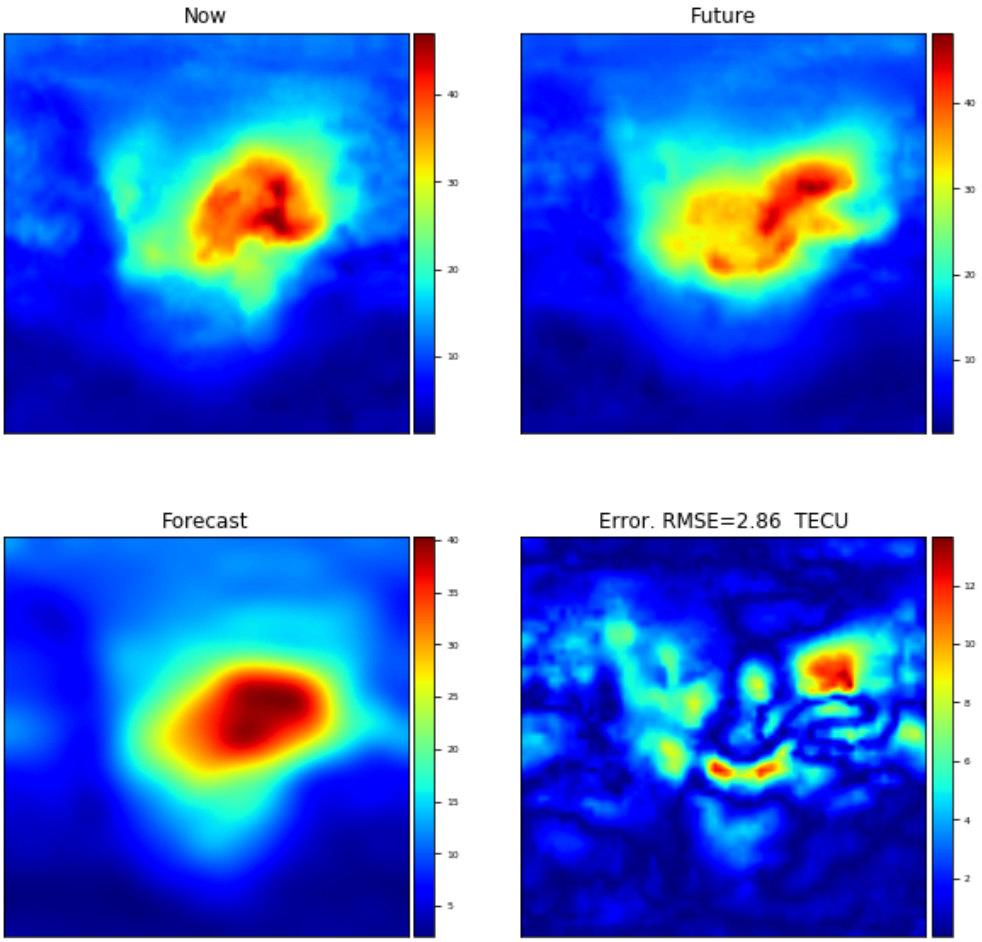

Figure 2. Forecasting by means of the DCT method at a horizon of 3 hours. Time stamp, 2016-02-21 13:15:00 (Year-Month-Day Hour:Minute:Second). Current map (upper left), future map (upper right), forecast (lower left), and forecast error (lower right).

In order to solve the problem of modeling simultaneously the smooth regions of the maps and the changes on the borders of the high-ionization regions we decided to use a different basis (in the sense of elements of a linear combination that span a subspace) that allowed for modeling the changes on the borders due to different distortions of the shape of the high-ionization regions. A set of transformations that might allow for modeling the changes at the borders, might be small rotations, small horizontal or vertical displacements, thickening or thinning of the borders, and hyperbolic (shear) distortions. The interest of these transformations resides in the fact that from one time stamp to the next, the changes on the map will be at the border of the ionized region and might be categorized as a mix of the above mentioned transformations.

A technique creating the basis that models the changes at the borders of the ionized regions is based on the idea of deformable prototypes. This technique is used in pattern recognition for dealing with common distortions in Optical Character Recognition (OCR) (see for instance, [28], [29] and [30]). The idea assumes that the image, in our case a map denoted as P, is a point in a high dimensional space. Each $71x72$ pixel map P is considered as a point in a space of dimension $71x72 = 5112$, i.e. $P \in R^{5112}$. For modeling the local changes from one map to the next, we will assume a transformation $s(P, \alpha)$, that creates another image, parameterized by α. In the exposition of the method, we will follow the notation as presented in [29].

In figure 3, we show a diagram of the method. Note that given a point P, we can create a trajectory in the space R^{5112}, by introducing small linear changes along directions in this space. In the example of

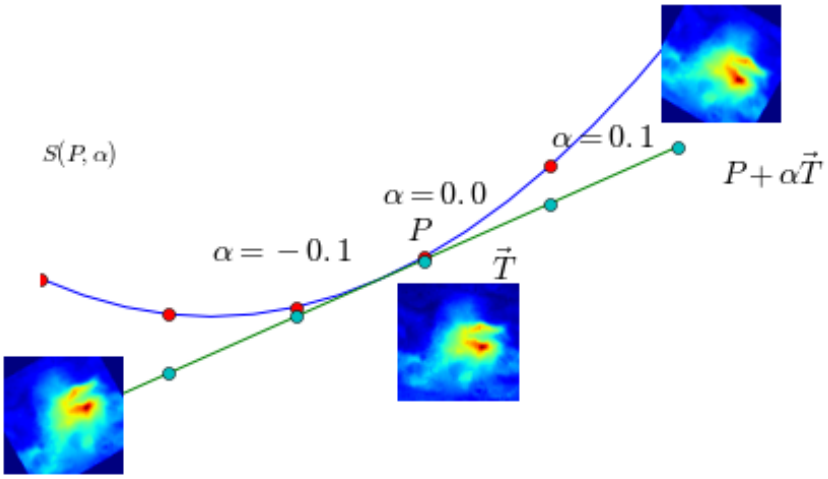

Figure 3. Tangent space to the Manifold of the map at '2016-01-19 10:30:00 (Year-Month-Day Hour:Minute:Second)'

the figure, the distortion, parameterized by α, consists on a rotation of the map. Thus, we will denote the set of points x in R^{5112} for each rotation α as the set $S_P = \{x | \exists \alpha \text{ for which } x = s(P,\alpha)\}$. In the figure, the trajectory followed by x, i.e. new images obtained from small rotations of P, is represented by a parabola, which is a one-dimensional curve embedded in the space of possible maps R^{5112}. Note that knowing the value of the pixels of the map P, and the structure of the transformation (in this case a rotation), we can characterize the trajectory in R^{5112} as a manifold, and one can compute the tangent vector at that point. The tangent vector is represented by a line that intersects the curved manifold at P, and is represented by the vector $\vec{T}$. For small values of α, we can approximate $s(P,\alpha)$ by the Taylor expansion around $\alpha = 0$:

$$s(P,\alpha) = s(P,0) + \alpha \frac{\partial s(P,\alpha)}{\partial \alpha} + O(\alpha^2) \approx P + \alpha\vec{T} \tag{1}$$

As an example, if we allow for another transformation, such as a vertical translation, we would have a different trajectory along a one dimensional manifold with the corresponding tangent vector. Hence, two tangent vectors: $\vec{T}_{rot}$ and $\vec{T}_{Y-trans}$, will define a tangent plane, and we can construct a first order approximation to a change on the map/point P by both transformations as a new point on the tangent plane:

$$s(P,\alpha) \approx P + \alpha_{rot}\vec{T}_{rot} + \alpha_{Y-trans}\vec{T}_{Y-trans} \tag{2}$$

Generalizing with the above example, we denote the estimate of the forecasted map at a future time value $\tau + h$ as $\hat{P}^{\tau+h}$, which is a function of the current map P^τ at time τ plus a point in the tangent plane spanned by the set of tangent vectors $\vec{T}^\tau_{dist}$. The tangent space $\vec{T}^\tau_{dist}$ is defined by the list of distortions/transformations $List_{dist}$ that we define in the following list:

$List_{dist}$ = { X/Y-translation, Thickening/Thinning, Rotation, Parallel Hyperbolic, Diagonal Hyperbolic }

The parameter h allows for modeling the fact that the forecasting horizon might be different from the sampling rate. The forecast at horizon $\tau + h$ can be expressed as follows,

$$\hat{P}^{\tau+h} = \alpha_0 P^\tau + \sum_{d \in List_{dist}} \alpha_d \vec{T}^\tau_d \tag{3}$$

A criterion for determining the value of the weights α_d of equation 3 might be the minimization of the $L2$ norm of the error between the observation at $\tau + h$, and the combination of the observation at τ, along with a linear combination of the components of the tangent space $\vec{T}^{\tau}_{dist}$, that is:

$$\{\alpha_d | d \in List_{dist}\} \leftarrow \underset{\substack{\alpha_d \\ d \in List_{dist}}}{argmin} \left\| P^{\tau+h} - \alpha_0 P^{\tau} + \sum_{d \in List_{dist}} \alpha_d \vec{T}^{\tau}_{d} \right\|^2 \quad (4)$$

Note that as the dimension of P is much higher than the number of parameters α_d, the problem is well posed. The initial model of Equation 4 can be extended and improved, taking into account the fact that the ionization regions change slowly in time, and therefore we could use a set of previous maps $\{P^{\tau-1}, \cdots, P^{\tau-N}\}$ in order to give a better forecast. The tangent spaces of the previous maps are also computed, and are combined linearly with the set of previous maps. This gives rise to two possible strategies for the estimation of the forecast:

a) For each past map $P^{\tau-i}$, estimate the values of α_d for each element of the tangent space by means of Equation 3, giving the partial estimate at $\tau + h$ from map i, $\hat{P}^{\tau+h}_i$. Afterwards the final estimation is done by a linear regression of each partial estimate.

b) Compute the tangent spaces of each of the maps $\{P^{\tau-1}, \cdots, P^{\tau-N}\}$, and do a linear regression using the N maps and the tangent space.

Although a linear combination followed by another a linear combination is equivalent to a single linear combination, note that from the estimation point of view, the coefficients obtained by method $a)$ are different from the ones obtained by $b)$. In the preliminary experiments (done using data of the year 2014), we have found that although the estimation method $a)$ allowed us to interpret the forecaster, and use tests of significance in order to asses the number of past maps relevant for the forecast. Nevertheless method $b)$ gave a better performance in the RMSE sense. An additional insight that we obtained from $a)$ was that the relevance of the previous snapshots for forecasting the future was not uniform along time, i.e., sometimes the most relevant snapshots were in the past few minutes, and in other cases the relevant snapshots were up to a few hours in the past.
For methodological reasons, we used the data from year 2014 as training for determining the structure and algorithms, and data from year 2015 as an independent validation and contrast dataset, and finally data from year 2016 to characterize the performance of the new forecasting model, i.e. the results are derived from data not seen when computing the parameters of the model and when deciding its structure. This methodology was done in order to diminish the possible effect of overfitting and selection bias. In the preliminary experiments we tested different machine learning techniques [32]. For selecting the forecasting algorithm we evaluated two experimental frameworks,

a) Use a historical set of values (one year) in order to estimate the parameters of the forecasting technique. This method has the advantage that can be used with a large set of machine learning algorithms. As a drawback, leads to complicated models that have to take into account all possible features of the time evolution of maps. Most of the machine learning algorithms can be trained with this approach. In particular we evaluated linear regression, LASSO, clustering (k-means), CART and neural networks (MLP).

b) Use only data from a limited number of maps prior to the current one. A drawback of this method is that it can only be used with regularized linear regression . This is because the other methods need a much higher number of examples. The advantage of this approach is that one can deal with the fact that, although the ionization pattern is similar over large regions, there is high variability of both, the shape, and morphology of the ionization regions in the time series of maps. The shape of the borders of the ionization regions changed slowly and the rate of change at different parts of the borders was not uniform. This justified both the introduction of the tangent space, that models the local changes, and the estimation of the coefficients of a linear regression with local data in order to select the relevant components.

The preliminary experiments showed that the method $a)$ produced a greater error than $b)$. We discarded other methods such as: neural networks, SVM, CART, because of the infeasibility of training the models with the approach $b)$.

As the evolution of the region of highest ionization was slow in the TEC forecasting case, and the changes between different time epochs were concentrated at the borders of the regions of highest ionization, we decided to use a method that took into account the small distortions. This was done by computing the tangent vectors (also called the Lie derivatives) for different kinds of locally spatial distortions. As each tangent vector in the space of dimension 5112 can be represented as an image of $71x72$, the forecast consisted of a linear combination of past maps and their respective set of tangent vectors for a set of distortions (for geometrical and algebraic details see next subsections).

3. Description of the Tangent Space method

In this section, following [29], we will define the equations for computing each of the the tangent spaces of the list of distortions $List_{dist}$ defined in section 2. A difference between the implementation for forecasting TEC maps and the one proposed in [29] for OCR, is that we did not smooth the finite difference approximation of the derivative. In the case of OCR, this smoothing is done by means of a 2D low pass filter and it is necessary in order to assure a smooth curvature of borders of the new template. In our case, we tested different smoothing techniques, such as linear low pass filters, and total variation smoothing, which resulted in a deterioration in the performance. Therefore, we approximated the derivative by a finite difference without smoothing: $\frac{\partial s(P,\alpha)}{\partial \alpha} \approx \frac{\delta s(P,\alpha)}{\delta \alpha}$

The finite difference is computed using second order central difference in the interior grid points of the map. At the boundaries of the map, the derivative was approximated by a first order difference between each border point (i.e TEC value) and the corresponding nearest inner grid point.

That is, for the interior grid points of the map, the discrete approximation was computed as follows:

$$\frac{\partial f(x,y)}{\partial x} \approx \frac{f(x+1,y) - f(x-1,y)}{2} \tag{5}$$

and

$$\frac{\partial f(x,y)}{\partial y} \approx \frac{f(x,y+1) - f(x,y-1)}{2} \tag{6}$$

For the border points of the map, the discrete approximation was computed as:

$$\frac{\partial f(x,y)}{\partial x} \approx \frac{f(x,y) - f(x-1,y)}{2} \tag{7}$$

and

$$\frac{\partial f(x,y)}{\partial y} \approx \frac{f(x,y) - f(x,y-1)}{2} \tag{8}$$

We construct the tangent vector for each of the possible seven transformations:

$List_{Transforms} = \{$ X-translation, Y-translation, Rotation, Parallel Hyperbolic transform, Diagonal Hyperbolic transform, Thickening and Thinning transform $\}$

We define all operations with reference the origin of coordinates at the center of the map, which in our case is: $(71/2, 72/2)$, and the local perturbation for each distortion is denoted by α. In figure 4, we show the tangent spaces for a given map (upper left) observed at time stamp '2016-02-21 13:30:00' (Year-Month-Day Hour:Minute:Second). The equations that define each of the directions that span the tangent space are the following:

1. X/Y-translation:

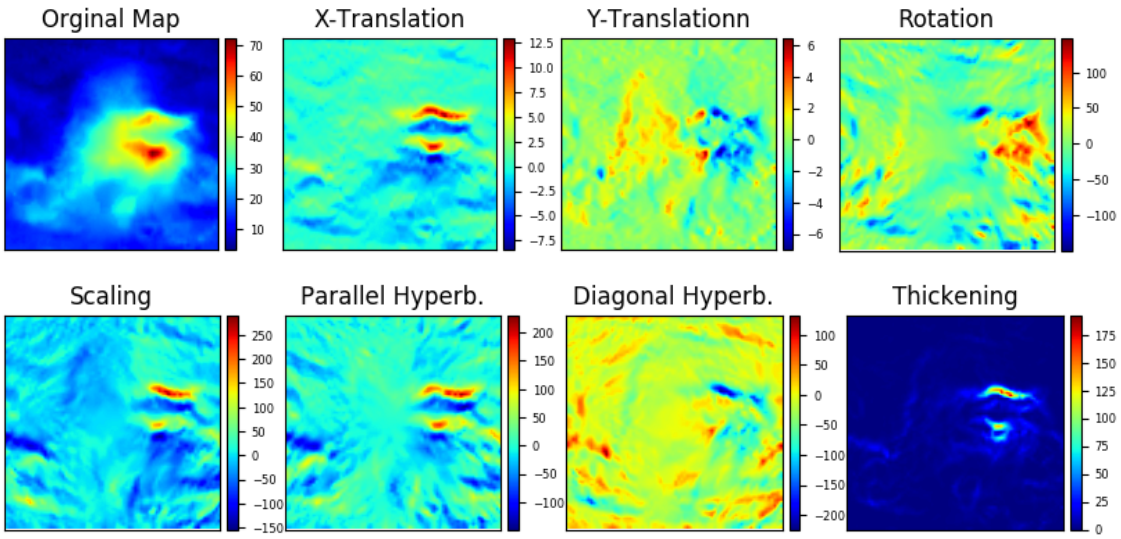

Figure 4. Components of the Tangent Space corresponding to a map observed at '2016-02-21 13:30:00'

$$t_\alpha^x : \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x+\alpha \\ y \end{bmatrix} \qquad t_\alpha^y : \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y+\alpha \end{bmatrix} \tag{9}$$

The corresponding Lie operators are defined as: $L_X = \frac{\partial}{\partial x}$ and $L_Y = \frac{\partial}{\partial y}$

2. Rotation by a small angle α:

$$t_\alpha^R : \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x\cos(\alpha) - y\sin(\alpha) \\ x\sin(\alpha) + y\cos(\alpha) \end{bmatrix} \tag{10}$$

The corresponding Lie operator is defined as: $L_R = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$

3. Parallel/Diagonal Hyperbolic transformation:
The parallel hyperbolic transformation defines a shear transform (left), and the diagonal hyperbolic transformation (right) defines a squeeze mapping.

$$t_\alpha^P : \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x+\alpha x \\ y-\alpha y \end{bmatrix} \qquad t_\alpha^D : \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x+\alpha y \\ y+\alpha x \end{bmatrix} \tag{11}$$

The corresponding Lie operators are defined as:
$L_P = x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}$ (parallel) and $L_D = y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$ (diagonal)

4. Thickening and Thinning transformation:

$$t_\alpha^T : \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x+\alpha r_x \\ y+\alpha r_y \end{bmatrix} \quad \text{where} \quad (r_x, r_y) = \frac{\nabla f(X)}{\|\nabla f(X)\|} \tag{12}$$

The corresponding Lie operator is defined as: $L_T = \|\nabla f(X)\| = \sqrt{(\frac{\partial}{\partial x})^2 + (\frac{\partial}{\partial y})^2}$

5. Scaling transformation:

$$t_{\alpha}^{S}: \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x + \alpha x \\ y + \alpha y \end{bmatrix} \tag{13}$$

The corresponding Lie operator is defined as: $L_P = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$

3.1. Transformation of a map along the tangent directions

The effect on a map of a small perturbation α in the direction of each of the components of the tangent space is summarized in figures 5, 6 and 7. Each figure presents two groups of maps, corresponding to two different directions in the tangent space. In Figure 5, we show the effect of moving the map, along the 'rotation' direction (left group), and along the 'Thickening' direction (right group). The left group of four maps show the original map (upper left), the tangent map for a small rotation of the map, and (lower) two reconstructed maps for opposite directions. The effect is moving the map either along the positive direction of the tangent space, i.e. $P + 0.1\vec{T}$ rotates slightly the map clockwise, or along the negative direction anticlockwise. Note that a shift of value $\alpha = 0.1$ in this case not only implies a noticeable rotation, but also increases the background level of ionization. Analogously, the tangent direction related with a thickening/thinning transformation (right), increases the area of the ionized region when added (thickening) $P + 0.1\vec{T}$, and reduces the area (thinning), when subtracted, i.e. $P - 0.1\vec{T}$. In Figure 6, we show the effect of a Scaling transform (left group), and a shift

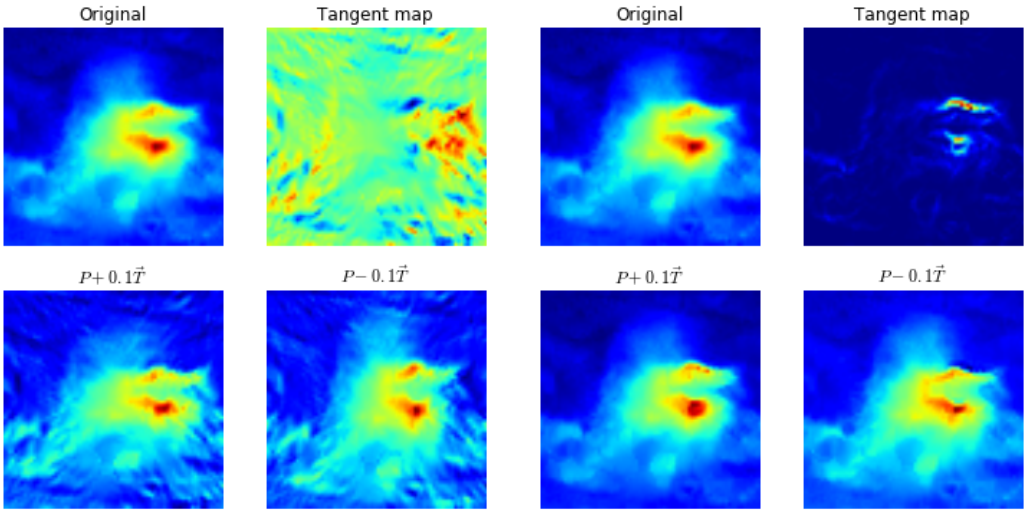

Figure 5. Rotation Manifold (left group) and Thickening Manifold (right group) at '2016-01-19 10:30:00'

in the vertical direction (right group). The effect of the scaling transform is a diagonal displacement of the ionized region, while the effect of the vertical shift is barely noticeable. This highlights that the size of the factor α can have different scales. In practice, as it can be seen for instance in Figure 9, the value of the weight of each dimension of the tangent space has different size, that is the value of α can take values in a wide range. While the terms related to rotation are small, the term related to the Y-translation is much higher. A value $\alpha = 0.1$ gives rise to a clear rotation of the ionized area in Figure 5, but the same value barely changes the map, as shown in Figure 6 (right). The case of a horizontal shift is analogous to the case of the vertical shift.

In Figure 7, we show the effect of the hyperbolic (shear) transforms of the map, which accounts for diagonal deformations of the map. Note that in this case, the scale of the factor α, increases the background ionization level, while it models reasonably well the changes in the highly ionized region. This effect of changing the background ionized level will be taken into account in the forecasting model by introducing a bias term in the forecast, $\alpha_{Intercept}$.

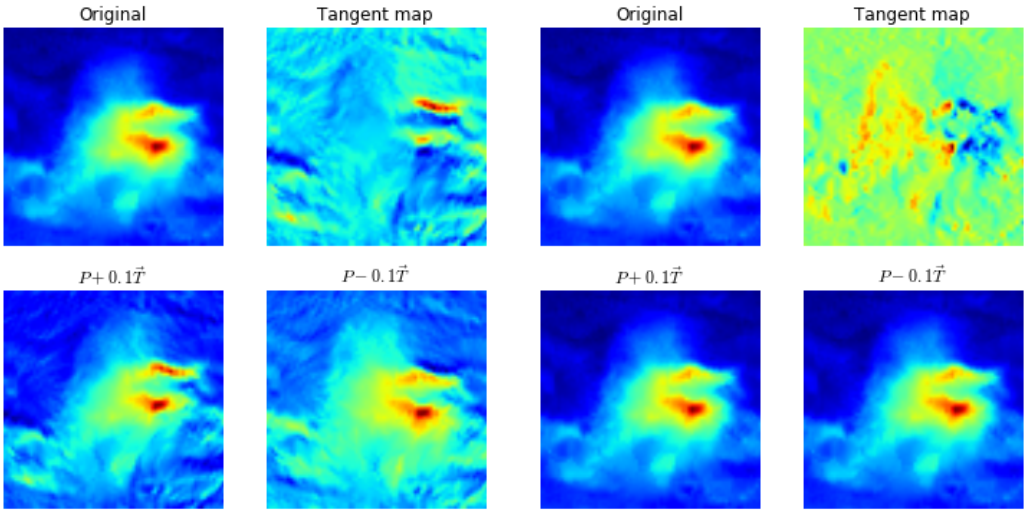

Figure 6. Scale transform (left group) and Y-translation (right group) at '2016-01-19 10:30:00'

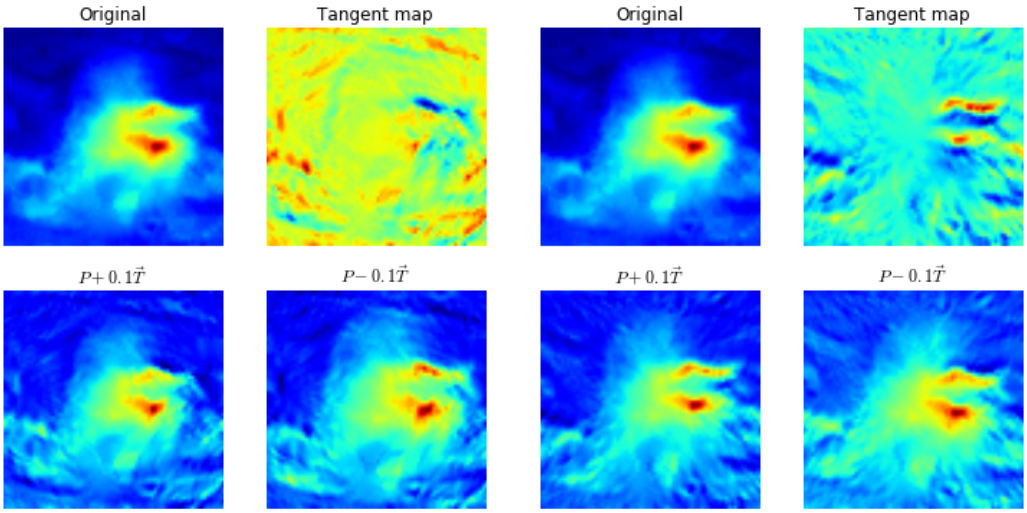

Figure 7. Diagonal Hyperbolic transform (left group) and Parallel Hyperbolic transform (right group) at '2016-01-19 10:30:00'

4. Model of the Forecaster

4.1. Structure

In this section, we present the model for the forecasting of the maps at a given horizon. This model is based on the ideas presented in section 2.

In univariate time series (see, for instance, [33]), the forecast model sometimes consists in a linear combination of the past samples. In our case, the analogy will consist in a linear combination of the time series of the last N maps $\{P^{\tau-1},\cdots,P^{\tau-N}\}$, with the corresponding tangent spaces. In the selection of the past lags, we took into account the fact that the general patterns of ionization have a main periodicity of about one day. Thus, along with the maps inside a time window near the current observation, we selected a time window at approximately a time lag of about 24 hours prior to the current map. The problem of model selection and model estimation is well known in machine learning

(see for instance [32], chapters 7 and 8). In the case of time series, it is a specially difficult problem, due to the correlation between samples and the presence of cyclical patterns of different origins and scales.

For methodological reasons, we divided the data into three sets in order to decouple the process of selecting the structure and the parameters of the model. For each potential configuration of the forecasting model (i.e. set of lags for previous maps and tangent spaces), we defined the optimal set of delays and estimation algorithm during the training year, i.e. 2014, and validated the performance for the different configurations with the RMSE evaluated on data in year 2015. Finally, independent data from year 2016 was left for reporting the performance. The list of time horizons we use in this work is the following: $\tau_{Horizon} = \{1/2h, 1h, 2h, 3h, 6h, 24h\}$

The structures for the forecasting that were evaluated included different time lags prior to P^{τ_0} , and time lags centered at multiples of 24 hours before P^{τ_0}. In addition, there was the possible inclusion of the tangent maps, with their possible combinations.

An interesting result is that the best performance was obtained when using as delay lags multiples of the forecast horizon, and the inclusion of a neighborhood of P^{τ_0-24h}.

The general structure of the models for forecast at horizon h included the use of the list of delay lags τ_i summarized at Table 1.

Table 1. List of time lags τ_i and weights α in the forecasting model. Note that the subindex 'd', runs along all the tangent spaces listed in $List_{dist}$

State	Output	Input 0	Input 1	Input 2	Input 3	Input 4	Input 5
Training	τ_0	$\tau_0 - \tau_h$	$\tau_0 - 2\tau_h$	$\tau_0 - 3\tau_h$	$\tau_0 - 24h$	$\tau_0 - \tau_h - 24h$	$\tau_0 - 2\tau_h - 24h$
Forecasting	$\tau_0 + \tau_h$	τ_0	$\tau_0 - \tau_h$	$\tau_0 - 2\tau_h$	$\tau_0 + \tau_h - 24h$	$\tau_0 - 24h$	$\tau_0 - \tau_h - 24h$
Weights		$\alpha_0, \alpha_{d,0}$	$\alpha_1, \alpha_{d,1}$	$\alpha_2, \alpha_{d,2}$	$\alpha_3, \alpha_{d,3}$	$\alpha_4, \alpha_{d,4}$	$\alpha_5, \alpha_{d,5}$

In Figure 9, we plot the values of the estimated parameters α, for a forecasting horizon of $h = 1hour$. This plot shows that the terms corresponding to the delay in the neighbourhood of P^{τ_0-24h} (i.e. one day) had a significant contribution. This effect was much lower when the delay lag was 48 hours, which we decided not to include because the use of this delay did not improve the performance on the validation database (i.e. year 2015), and increased significantly the complexity of the model. Another effect that consistently appeared in most of the experiments was that most of the contribution to the forecast came from the current map, along with its tangent space, and the forecasting power of greater lags diminished quickly. Thus, the context for forecast is limited to $\tau_0 - 2\tau_h$. Also in order to model the fact that there might be a global change of the ionization, we introduced an intercept $\alpha_{Intercept}$. The forecasting was done by means of the following formula,

$$\hat{P}^{\tau+h} = \alpha_{Intercept} + \sum_{L \in List_{Lags}} \alpha_L P^L + \sum_{d \in List_{dist}} \sum_{L \in List_{Lags}} \alpha_{d,L} \vec{T}_d^L \tag{14}$$

The set of elements of the tangent space are summarized in $List_{dist}$, and explained in detail at the subsection 3.

4.2. Parameter Estimation

In this subsection we will argue the criteria to select of the estimation algorithm for the model presented in equation 14. The model consists of a linear combination of maps, weighted by the coefficients α_L and $\alpha_{d,L}$, which can be estimated with different criteria. That is we will set a structure of delays, and the values of the coefficients α, will be determined at each forecasting step by minimizing a cost function, and therefore making the forecast robust with respect to the solar maximum/minimum conditions. In analogy to equation 4, we will minimize the forecast error with some restrictions on the possible weights. Note that the estimation of the parameters is done using only the near past of the map P^{τ_0} at time τ_0. Specifically, in order to estimate the parameters of the model we used as

time indices the row 'Training' at Table 1, and the forecast was done using the time indices at the 'Forecasting' row. Note that in the same table, in the lower row, we define the notation of the weights associated with each delay.

The use of only the last few samples (i.e. just the samples in the row 'Training' of Table 1) for the estimation of the weights is justified by the empirical fact that the performance of the forecaster, exhibited a better RMSE. When the estimation of the parameters (i.e. the training) is done using the whole year the resulting performance was significantly worse. The explanation is that the local (in the temporal sense) variation of the maps, has a greater contribution to the relative weighting of the past samples and tangent space than a fixed set of weight computed from a long time series.

A drawback of using a small set of samples for the estimation of the parameters is that the system of equations might be underdetermined and will suffer of problems of collinearity. Therefore in order to deal with this problem, the criterion for the estimation of the parameters will have a regularization term. This regularization term consists on a penalization of the norm of the weights. In case of using norm $L2$ the method is known as Ridge Regression, and the case of using norm $L1$ the method is known as LASSO (see for instance [32] chapter 3). The difference between both norms is that the Ridge Regression introduces a constant λ in the diagonal of the covariance matrix which results in a uniform decrease in the value of the estimated parameters, while in the LASSO algorithm the parameter λ determines the level of sparsity, i.e. the number of parameters with a value equal to zero.

The optimization criterion is presented in equation 15, where the regularization term $\left\|\alpha_d, \alpha_{d,L}\right\|_{Li}$ corresponds to Ridge Regression when $Li = L2$, and LASSO when $Li = L1$. We also considered the use of norm $L1$ in the error term, i.e. $\left\|P^{\tau+h} - \hat{P}_i^{\tau+h}\right\|_{L1}$. This norm penalizes small values of the error, and the influence of the outliers on the estimated weights is small. In our problem at hand, the use of this norm on the error term, gave a much higher proportion of maps with at least a negative values of TEC. Therefore we selected $L2$ as norm for the error term, giving the following cost function. In the implementation done in this work, the results are presented using the Ridge Regression regularization, and the λ was selected by cross-validation on the data from the year 2014. The difference in performance obtained from using Ridge Regression or LASSO was negligible, and the Ridge Regression was selected because the running time was faster.

$$\{\alpha_d, \alpha_{d,L}\} \leftarrow \underset{\alpha_d, \alpha_{d,L}}{argmin} \left\|P^{\tau+h} - \hat{P}_i^{\tau+h}\right\|_{L2} + \lambda \left\|\alpha_d, \alpha_{d,L}\right\|_{Li} \tag{15}$$

A typical example is shown in Figure 8 we present the forecast at a horizon of 3 hours, at two different time stamps, i.e. at 2016-05-12 15:15:00 (left) and 2016-02-21 13:15:00 (right). For comparison purposes in Figures 1 and 2 we show an example of the performance when using a DCT forecast in the same conditions. The benchmark that we will follow for determining the performance of the system will be the use frozen maps, which we define as the prediction made with the current map keeping its VTEC values constant in a Sun-fixed reference frame, i.e. considering local-time / latitude coordinates. The RMSE ratio between the forecast, and using a frozen map as forecast was of 76%, which as can be seen in Figure 12 is about the mode of the histogram of ratios of the RMSE. The weights associated with the estimate are shown in Figure 9. Note the presence of a positive intercept or bias $\alpha_{Intercept}$, which models both, the contribution of the background ionization of the different elements that are combined linearly, and that the possible increase of the background ionization. In addition note that the main contribution to the forecast, come from the current map, with a strong negative component of a vertical deviation. The third most important contribution to the forecast is related to the map at $\tau_0 + \tau_h - 24h$, with it's tangent space long the vertical direction. On the other hand the contribution of the maps at 1 and 2 hours in this case is marginal.

The forecast map $Vect_{Pred}$ (lower right) is much smoother than either the current map P_{now} or the target map P_{Future}. In addition the distribution of the error reproduces the observation that the borders of the ionized region accumulate most of the error, in this case in part of the border the error reaches a maximum of 9.5 TECUs.

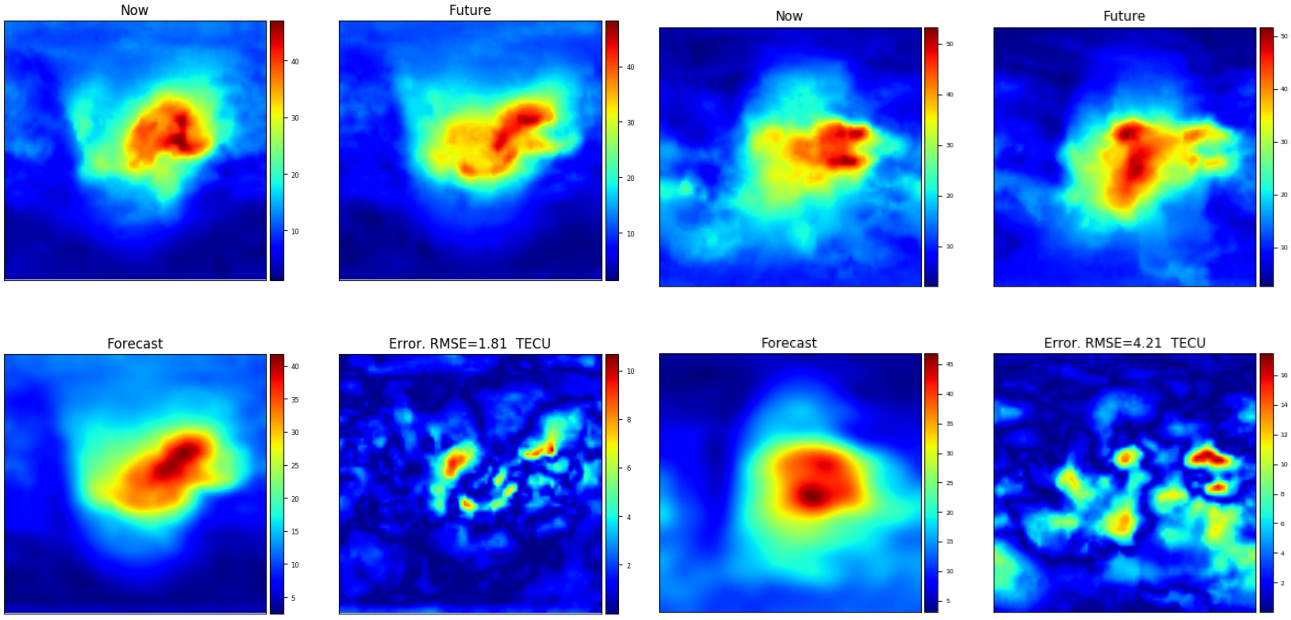

Figure 8. Forecasting at a horizon of 3 hours (Tangent space method), at time stamps, 2016-05-12 15:15:00(left group) and 2016-02-21 13:15:00 (right group). For each group we plot, current map (P_{now} upper left), target map (P_{Future}, upper right), Prediction error (lower left), and Forecast ($Vect_{Pred}$ lower right).

The forecasted TEC values were always positive (in the train and validation years), except for a small fraction of cases in the 30 minutes and 1 hour forecast. In the case of the forecast at 1hour, there were only 123 maps that had at least one negative pixel, of a total of 34536 (i.e a whole year). In this case, the negative values, can be substituted by the values of the Frozen prediction at the latitude/longitude where the forecast was negative.

5. Results

In this section we will summarize the performance of the forecast system based on the tangent spaces. The first results consist in the comparison of the error in the sense of total TEC RMSE (i.e. the mean value for all the pixels of the map). We compare the forecast by using tangent spaces (which we denote as tangent space method) with a forecast created considering no changes in the map in a mostly Sun-fixed reference frame (local-time vs latitude),which we denote as frozen map method.

Given a time series of N_τ maps, for the map at time τ consisting of the TECU value $x^\tau_{lon,lat}$ and the estimated values $\hat{x}^\tau_{lon,lat}$ the RMSE was computed as:

$$RMSE = \sqrt{\frac{1}{N_{lat} \times N_{lon}} \sum_{lat=1}^{N_{lat}} \sum_{lon=1}^{N_{lon}} (x^\tau_{lon,lat} - \hat{x}^\tau_{lon,lat})^2} \tag{16}$$

In Figure 10 we shown a comparison of the forecast RMSE time series for the tangent space method with the frozen map method during an interval of 10 days during the month of March 2016. In the figure the comparison between methods has been done for horizons of 30 minutes, 1 hour, 3 hours and 1 day. The RMSE time series for the tangent space method (red), is systematically below the frozen method (blue) for the cases of 30 min, 1 hour and 3 hours, with periodicity of the error of about a day. In the case of the horizon at 1 day, the forecasting error is the same for both methods. The RMSE results are computed as a mean over all the map, and as mentioned in section 2, the main contribution to the error originates from the borders of highly ionized regions. Therefore the TEC values presented in this subsection are an average between low error regions and these borders. The

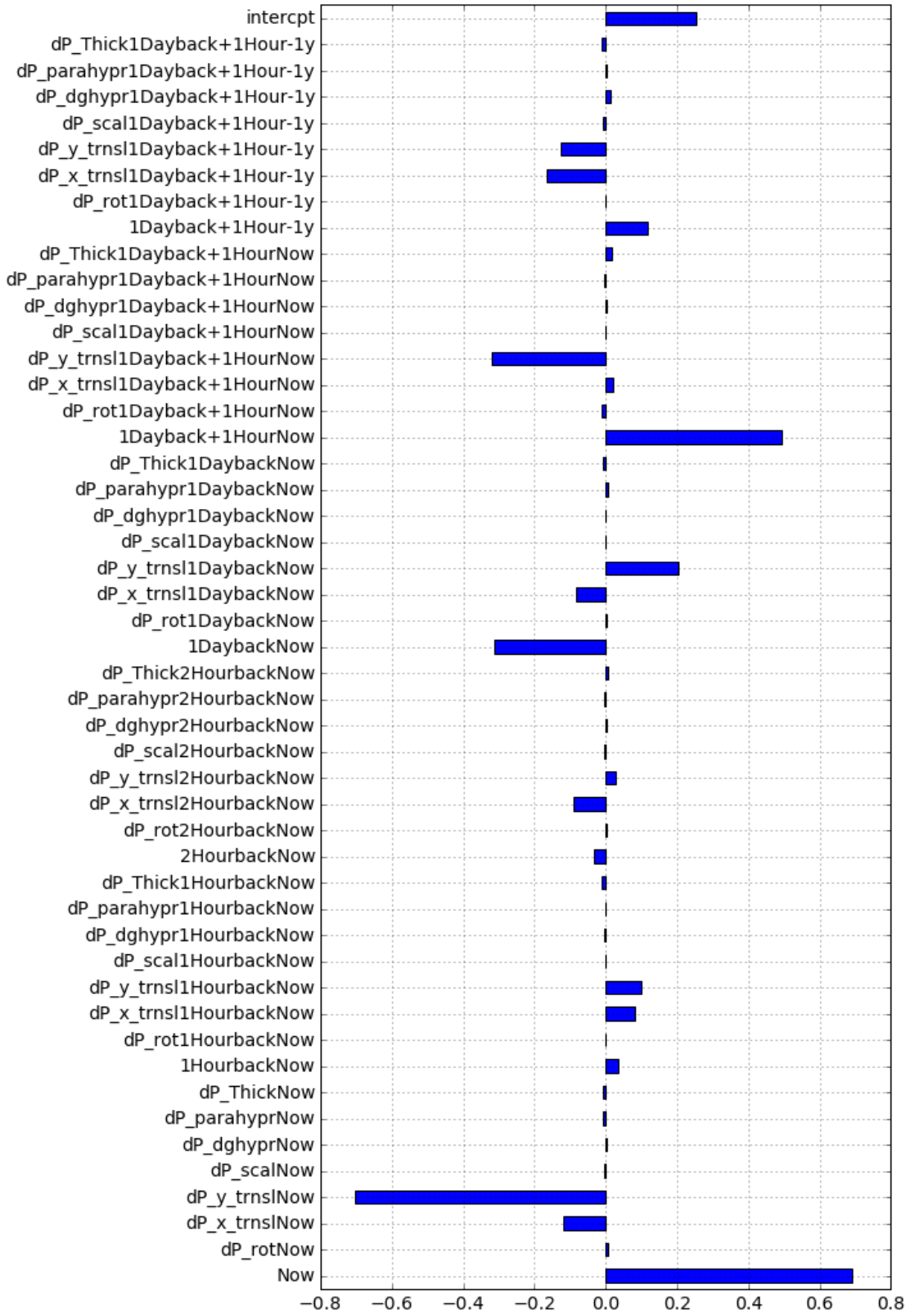

Figure 9. Weights for a one hour forecast at time '2016-02-21 13:15:00'. The term $dP_{distTime}$, indicates the tangent space $\frac{\partial s(P,\alpha_{dist})}{\partial \alpha_{dist}} = \vec{T}^{\tau}_{dist}$ and the corresponding time lag.

result is representative of the general behaviour of the algorithm, and is compatible with the results averaged over an entire year, as shown in Table 2.

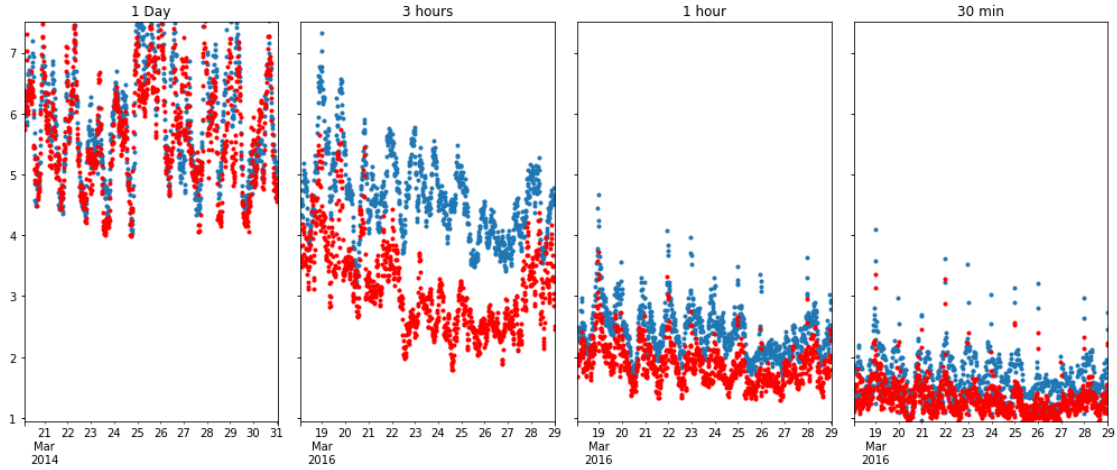

Figure 10. Time evolution of the TEC RMSE, for the case of forecast with the tangent space method (red), and frozen map method (blue) for different horizons (see annotation on the upper part of the figures)

In Figure 11 we present the boxplots (see [34]) of the forecast errors for different horizons, computed for a whole year. One can see that in all cases, except for 1 day, the forecasting error in terms of forecast RMSE obtained by tangent space method was better than using a frozen map. The proportion between of improvement is summarized in Table 2. The improvement not only refers to the median value, but also to the range of the errors between the external whiskers (defined as the range between the quartile Q1 (i.e. the percentile 25%) less 1.5 times the interquartile range, i.e. the difference between Q3 (i.e. the percentile 25%) and Q1, that is, the first datum greater than $Q1 - 1.5 * (Q3 - Q1)$ and the other extreme is given by the last datum less than $Q3 + 1.5 * IQR$). As can be seen in the figures, the dispersion around the mean corresponding to the tangent space method is lower, in the senses that encloses a smaller margin, and a also in the sense that systematically has a lower value. Another feature is the lower prevalence of high values of the outliers in the case of the tangent space method.

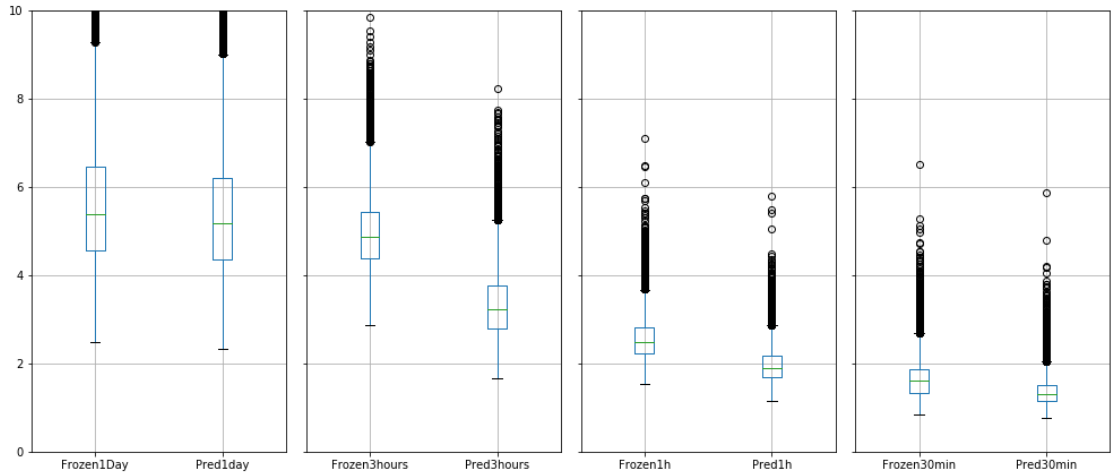

Figure 11. Box plot of the TEC RMSE, for the case of forecast with frozen value (left) and the tangent space (right), for each horizon (see the tick labels of each figure)

Table 2. Ratio of the RMSE between tangent space forecast and frozen maps wrt actual maps for different horizons.

Horizon	1/2h	1h	2h	3h	6h	24h
Ratio	84.99 %	77.65%	71.35%	69.34 %	87.23 %	95.76%

In Table 2, we show the ratio of the root mean square error (RMSE), of the tangent space method vs the frozen map method. The RMSE in this case is computed as a mean over the N_τ, maps corresponding to a year, where each forecast is made every 15 minutes. The RMSE was computed as follows:

$$RMSE = \sqrt{\frac{1}{N_\tau \times N_{lat} \times N_{lon}} \sum_{\tau=1}^{N_\tau} \sum_{lat=1}^{N_{lat}} \sum_{lon=1}^{N_{lon}} (x^\tau_{lon,lat} - \hat{x}^\tau_{lon,lat})^2} \tag{17}$$

From the table it can be seen that in all cases the ratio of forecast RMSE is below 100%, that is the tangent space method gives a lower RMSE forecast in mean. In Figure 12 we show the histograms of the ratio for a time interval of 42 days. Note that the maxima of the histograms coincide approximately with the mean values shown in Table 2, which is due to the low variance and skewness of the histograms. Note also that as the coefficients of the tangent space forecaster are computed for each forecast, the results performance of the method does is not biased by the solar cycle, that is, we have seen empirically that the total RMSE is proportional to the underlying mean TEC. Thus although the tangent space method is better than freezing, the RMSE will also depend on the solar cycle conditions.

The results for different horizons show a performance that follows a convex shape, with a minimum error at a 3 hours horizon. In particular the performance at 1/2 hour and at 6 hours, has a mean value of about 85%, which worse than at horizons of $\{1, 2, 3\} hours$. In the case of 1/2 hour this is due to the fact that the rate of change of the state of the ionosphere is slow, and at this horizon there is small room for improvement in the forecast. This small change of the state of the ionosphere is reflected in the fact that the RMSE due to freezing is in mean about 2 TEC, which is lower than at other forecasting horizons as seen in Figure 10. Interestingly in the histogram shown in Figure 12, the number of samples with error greater than 100% is much higher than in the other cases and also the upper values are much higher, this is accounted by the fact that the histogram is a ratio between numbers that can be small.

A trend that can be seen in Figure 12 and in the boxplots, is that the spread of the forecast errors errors increases as a function of the horizon. Nevertheless note that as reflected in Figure 12, the tangent space forecast for intermediate horizons gives a lower limit of the % of error, and the number of samples for an error greater than 100% is much lower. Which indicates that the quality of the tangent space forecaster in the sense of total RMSE is better than in the case of freezing, in all the horizons except 24 hours.

On the other hand, at a horizon of 6 hours the performance degrades, being comparable to the performance at 1/2 hour, which indicates that the hypothesis of small changes in the tangent space (see section 3.1), begins to not be entirely valid. And Finally at a horizon of 24 hours, the performance does not differ much from freezing the maps, which indicates both, that the hypothesis underlying the model of tangent spaces does not hold, and that also that a simple linear combination of past observations is not enough for modeling a TEC map at a day's horizon. The forecast by means of the tangent space is slightly better than just freezing the last observation, Figure 12 shows that the ratio between the performances of the two systems has a shape similar to a Gaussian, with a mean at 95%, with a standard deviation nearly 10%. On the other hand the boxplots in Figure 11 (left), show a median RMSE of 5 TECUs, which might not be acceptable. Note that although we have found that the component at 24 hours is relevant for the forecast at short term horizons, it does not seem adequate

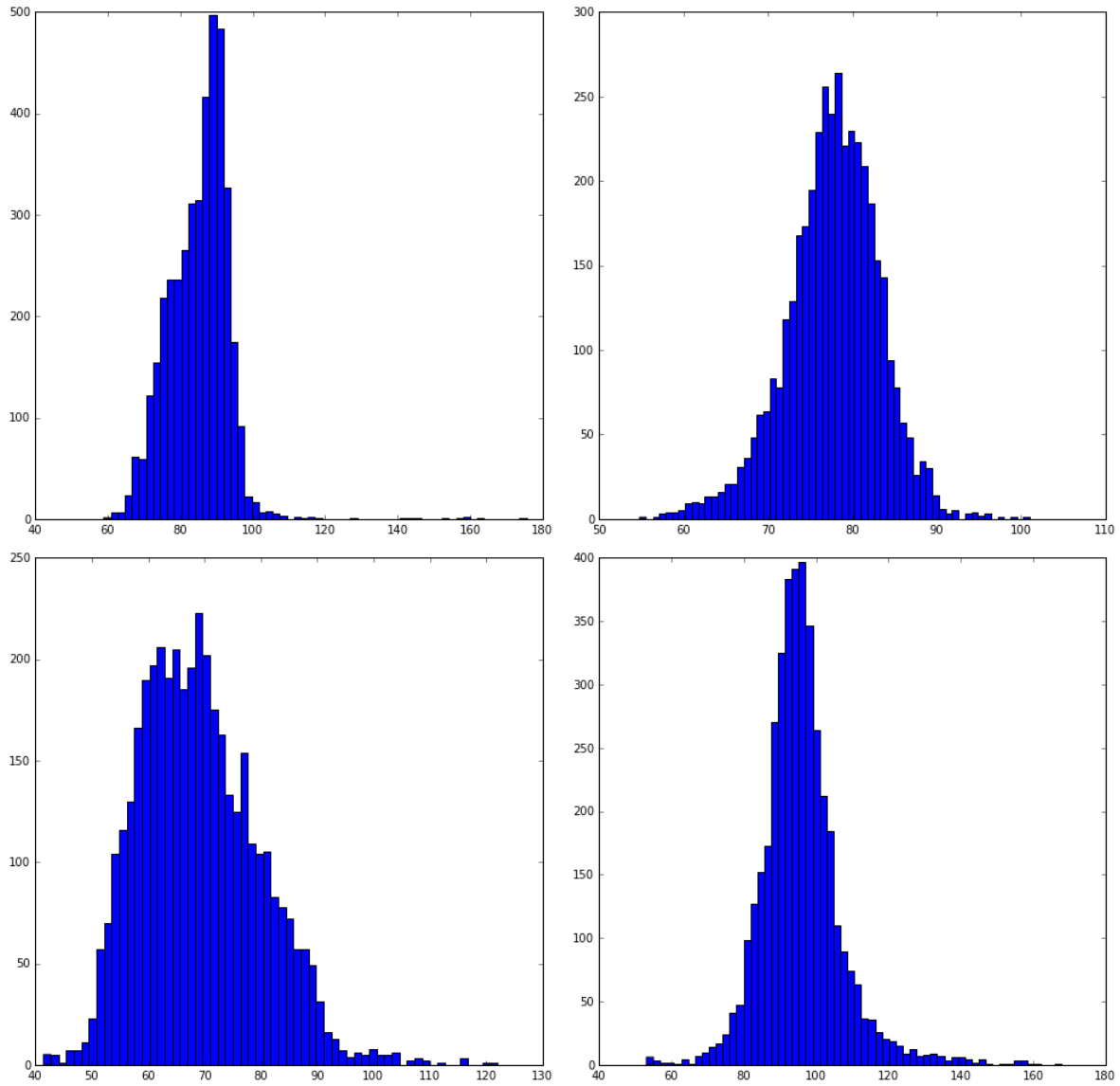

Figure 12. Histogram of the forecast RMSE ratio between the tangent space method, and a frozen map method. Upper left horizon at 30 min, upper right horizon at 1h lower left horizon at 3h and lower right horizon at 24h. From 2016-01-10 to 2016-02-21, a total of 4080 maps.

for a forecasting at such a long horizon. That is, for short-term predictions, the maps at 24 hours in advance complements the information from the last observation and helps to improve the forecast by making use of the periodicity of the ionization patters. On the other hand, when the horizon is directly 24 hours, the information yielded by the last observation is not enough for a good forecast.

For comparison purposes, in Figure 13, we show the time evolution of the RMSE in TECUs for a whole year and a different forecast horizon (3 hours). The figure consists on the superposition of the instantaneous RMSE and the low pass filtered (zero phase, moving average with a cut off frequency of $1/(0.5month)$) time series. The shape of the temporal evolution of the RMSE over a whole year is similar for the other forecast horizons. The first feature that we notice is a seasonal component with a maximum activity at spring and autumn and a superposed a cyclic component with a period of 28 days. In the case of the forecast by means of the tangent space model, the difference between the maximum and minimum of the low pass filtered version is of 1.5 TECUs, with a mean value in spring of about 3.5 TECUs, in summer of 2 and in winter of 2.5 TECUs. The variability of the error is much

lower in summer. The variability is much higher than would be expected in the case of a Gaussian distribution (see Section 5.4), with local peaks over 6 TECUs occurring only certain moments in 14 days of the time series. A test was done with the years 2014 – 2015, where the activity was higher, and the results were analogous, with a systematic increase of the low pass filtered RMSE of 1 TECUs. Note that, from a methodological point of view (of machine learning), this analogy is merely indicative, because these years were used for determining the structure of the forecaster.

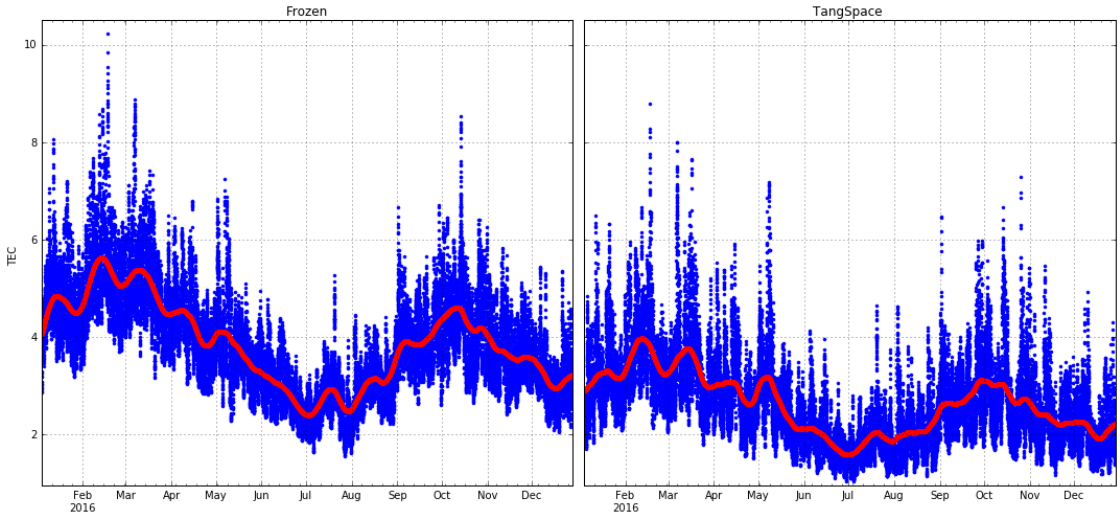

Figure 13. Time evolution during a whole year of the total RMSE in TECUs for a horizon of 3 hours. Left, the frozen forecast, in red the time series low pass filtered. Right the forcast by means of the tangent space method.

5.1. Improvement on the forecast due to the use of the tangent space and comparison with the DCT forecasting method

In this section, we compare the proposed approach with a forecasting system that does not use tangent spaces but just the original VTEC maps (labelled Maps Only) and with one based on the DCT (DCT method). In both cases, the number of free parameters of these models is lower than that in the tangent model, which means that the comparison is not in equal terms. On the other hand, a possibility of doing this comparison with an equal number of degrees of freedom would be to use a number of maps equal to the number of elements in the tangent map. In this case, either if we took maps further in the past, which are increasingly different from the map to be forecasted, or if we increased the number of recent maps, highly correlated inputs are given and therefore we have a badly conditioned problem. The improvement due to the use of the tangent spaces, is shown in Table 3. The result is that the use of the tangent space coefficients improves the forecast result in all horizons, except for 24h. In all cases (except for the horizon of 24h) the use of tangent space improves the results to the DCT method. The 24h case is explained by the fact that the tangent spaces model local distortions, which do not account for the changes at such a long range. Note that the confidence margins (standard deviations) are not shown, because using 34626 maps for computing the mean values of the RMSE gave extremely small values.

5.2. Performance as a function of the latitude

Next we analyze the performance of the algorithm as a function of the latitude. We analyzed the whole year 2016 and we set a horizon of 1 hour, with a total of 34626 maps. Note that the test year is different from the years used for determining and validating the structure of the forecaster. For

Table 3. Improvement Ratio of the RMSE between tangent space forecast and frozen maps when using the tangent space forecaster vs the case of not using the tangent space of each of the maps.

Horizon	1/2h	1h	2h	3h	6h	24h
Ratio Tang Space method	84.99 %	77.65 %	71.35%	69.34 %	87.23 %	95.76%
Ratio using Maps only	89.08 %	80.90 %	76.95 %	73.07 %	91.45 %	94.08%
Ratio DCT Method	87.75 %	78.25 %	74.38 %	71.21 %	88.94 %	89.11%

presenting the results we have selected a horizon of 1 hour, in order to show the performance in a situation between the best and the worst case.

The RMSE as a function of the latitude is defined as;

$$RMSE(lat) = \sqrt{\frac{1}{N_\tau \times N_{lon}} \sum_{\tau=1}^{N_\tau} \sum_{lon=1}^{N_{lon}} (x^\tau_{lon,lat} - \hat{x}^\tau_{lon,lat})^2} \tag{18}$$

where $x^\tau_{lon,lat}$ at the begining of Section 5. In Figure 14, we can compare, for each latitude the mean value of the RMSE in TECUs of the tangent space method vs. the frozen map method. The results of the tangent space method (left) systematically give a lower value of RMS compared with the frozen map method (right) for all the latitudes, with the most significant improvement between 30 degrees north and 30 degrees south. With a maximum difference of about 1 TECU for the both maxima, i.e. 8 degrees north, and 20 degrees south. On the other hand for latitudes over 50 degrees (north or south), the performance of both methods is similar, due to the fact that ionization activity at these latitudes is much lower than near the equator, and in both cases the RMSE is about 1 TECU. At the equator there is a dip in the RMSE profile. Here the tangent distance forecast is better by 0.5 TECUs. The fact that most of the forecast error is located around ± 30 degrees and the shape of the profile, can be explained from the spacial distribution of the forecast error, that tends to be located at the margins of the ionized zone as was shown in Figure 8. The distribution of the error on the map also explains the dip on the forecasting error at the equator. In Figure 15 we show the profiles of the Root Mean Square of the values in TECUs, for the whole year 2016 (left) and the ratio of RMSE the tangent space method to the frozen map method (right) also for the year 2016. In both cases we present the results as a function of the latitude. In In Figure 15 (left), the mean TECU values are concentrated around the equator, with a mean value of for the case of latitudes near the equator of about 25 TECUs, and decreasing as the latitudes get near the poles. Note that a first comparison with Figure 14, tells us that the forecasting error in mean is about 10%. This last fact, should be nuanced with the fact that the forecast error is not uniform on the map, but distributed on the borders of the ionized part of the map. In Figure 15 (right) the ratio of the tangent distance forecast to the frozen map, shows two minima, located at (9-20) degrees north and (13-25) degrees south, which is about 75%, which indicates that the use of the tangent distance forecast produces a reduction of a 25% in mean on the forecast error, in the regions of higher ionization. That is the best improvement, is given at two crests of the Appleton-Hartree anomaly, side by side of the equator. Therefore the tangent distance method is capable of modelling regions with high TEC gradients. Note that as the coefficients α of the forecaster are re-estimated for each prediction, which means that the movements of the anomaly can be tracked on real time. The forecast improvement at the regions near the pole is lower, but this can be explained because these are regions of lower activity and show a smaller rate of change. This figure hints at the possibility of designing a performance index that might assign a different penalization to the errors depending on the level of activity of related to the latitude.

5.3. Bias Variance decomposition of the forecast error

Next we analyze the Bias Variance decomposition of the forecast error, which we define as follows: For given a map P and the forecasted map $\hat{P}$, the error can be decomposed as:

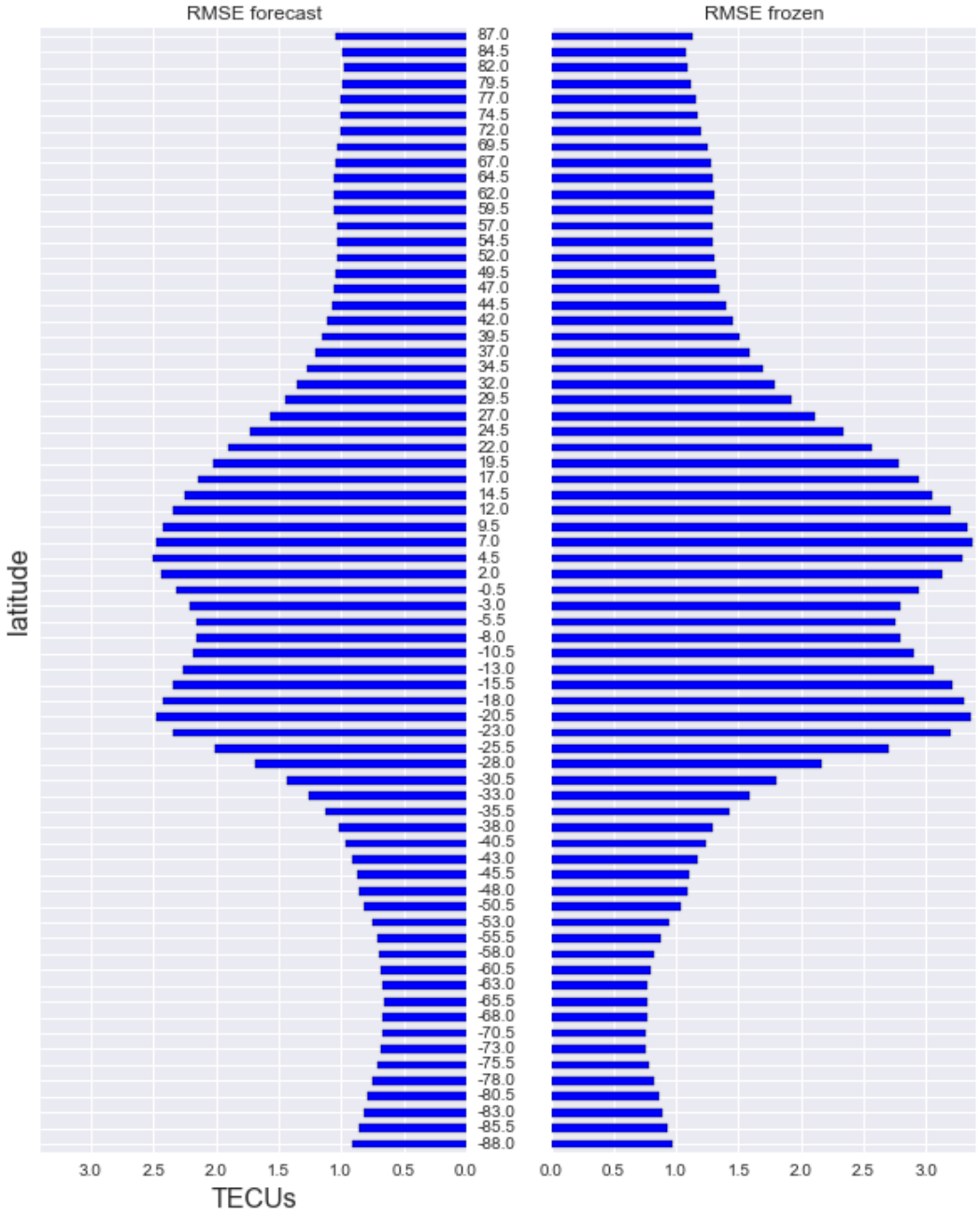

Figure 14. RMSE of tangent space method (left) vs. RMSE of frozen map method (right) as a function of the latitude.

$$E\left[\left(P-\hat{P}\right)^2\right] = Bias[\hat{P}]^2 + Var[\hat{P}] \tag{19}$$

where each term is defined as:

$$Bias[\hat{P}] = E\left[P-\hat{P}\right] \tag{20}$$

and

$$Var[\hat{P}] = E\left[\hat{P}^2\right] - E\left[\hat{P}-P\right]^2 \tag{21}$$

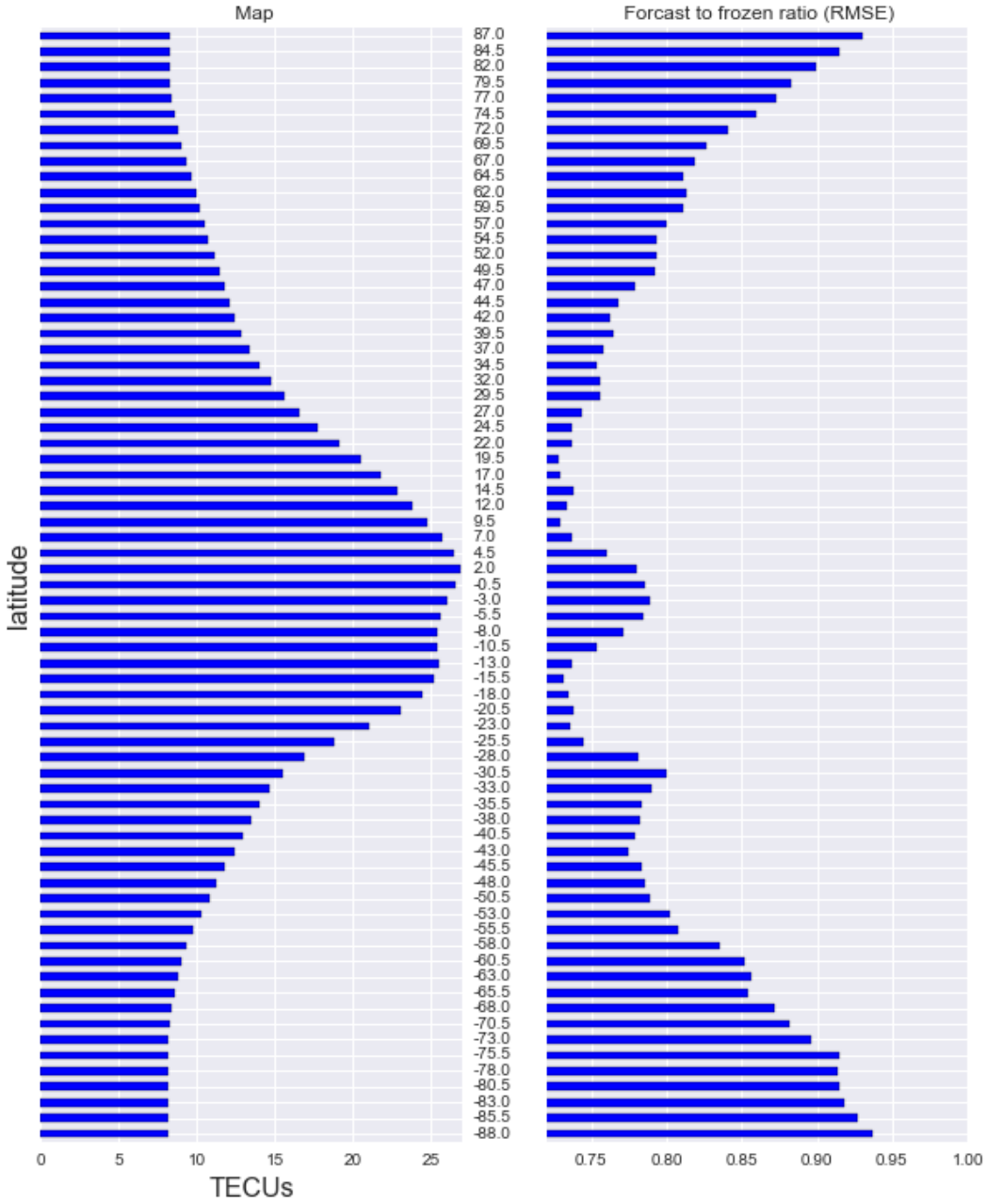

Figure 15. Root Mean Square in TECUs as function of the latitude (left). Ratio of RMSE the Tangent Distance Forecast Error to the Frozen forecast error (right).

Figure 16, shows the analysis of the contribution to the error of the $Var[\hat{P}]$ (left in $TECUs^2$) and the $Bias[\hat{P}]^2$ (right in $TECUs^2$) as a function of the latitude. The variance of the forecast near the equator is of about 12 TECUs, which can be explained by the fact that the main contribution to the error appear concentrated at the borders of the ionized regions, and the variation of this moves in this regions. As it can be seen for instance in Figure 8, the high error values occupy a small part of the map, localized at the edges of the high ionization region. The value of these peaks can reach the order of 9 TECUS, while the error in the inner parts is much lower, of the order of 1 TECU. Therefore, the observed variance should not be understood as the variance of a Gaussian distribution, but as the variance of a bimodal distribution.

On the other hand, the forecast system is almost unbiased (see Figure 16 right), the contribution of the bias to the total error is always less than 0.1%. Note that the figures have a different scale. The explanation of the low bias of the tangent distance forecast error is associated with the most effective set of parameters describing the degrees of freedom of the system. Note that in Section 4.2, we comment how we dealt with the fact that the estimation of the parameters α might be undetermined.

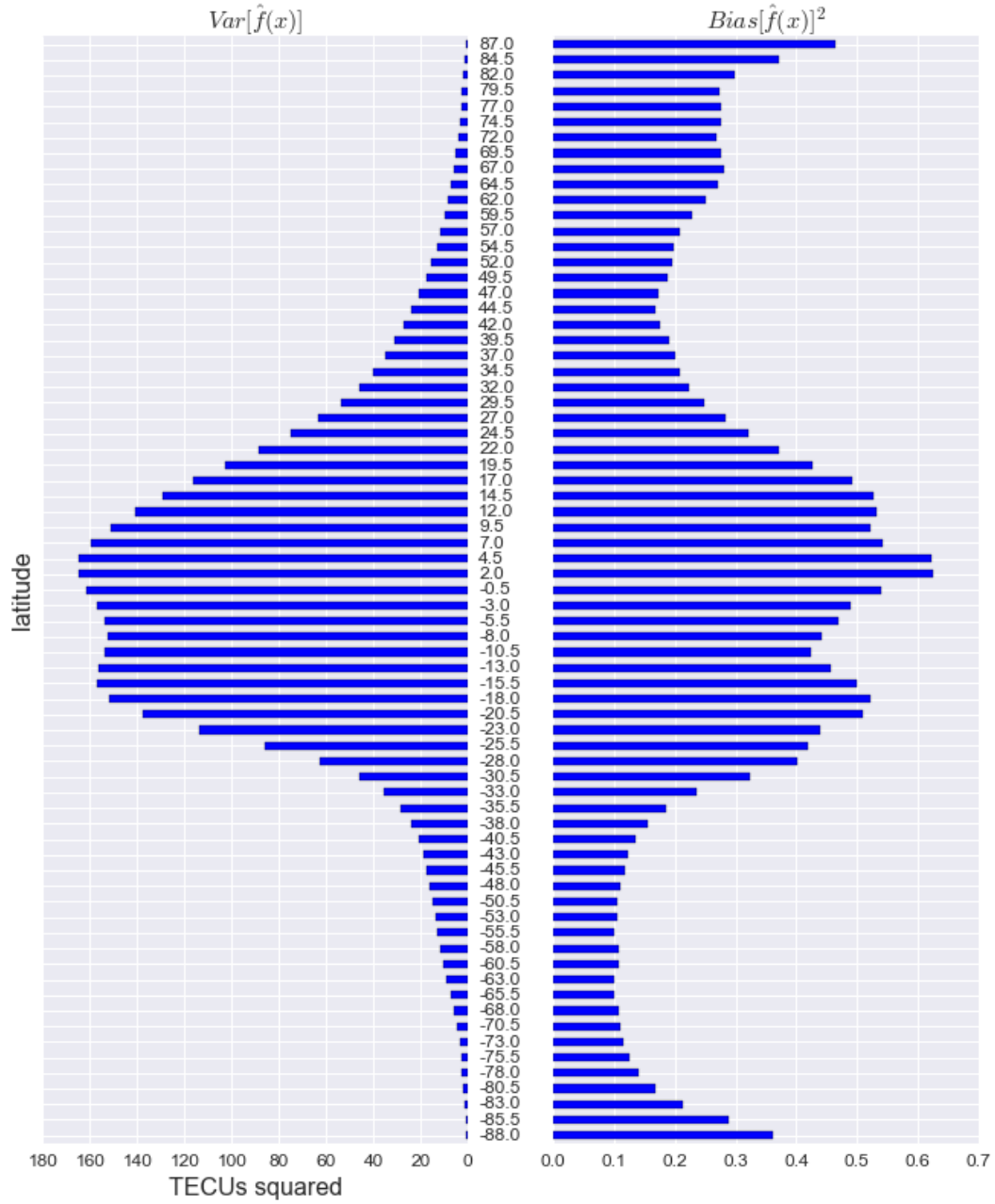

Figure 16. Decomposition of the forecast error variance '$Var[\hat{f}(x)]$ (left) vs Bias $Bias[\hat{f}(x)]^2$ (right). Note that the units are $TECUs^2$

5.4. Histogram of the forecast error

The histogram of the forecast error in TECUs is shown in Figure 17. The histogram is computed for all the pixels of the 34626 maps, which gives a number of samples greater than 166 million, that is the size of the map times the number of maps,i.e. $72x72x34626$. In particular, the high error values are

Table 4. Empirical fraction of values greater than a TECU value .

TECU	2	4	6	8	10	12
Prob (τ > TECU)	1.95×10^{-01}	1.48×10^{-02}	9.84×10^{-04}	1.21×10^{-04}	2.81×10^{-05}	8.16×10^{-06}

located at the borders of the highly ionized regions, and account for the tails of the error histogram. In the log frequency histogram we can see that the distribution follows approximately a line, and can be described by :

$log_{10}(TECU = t) = -1.27t + 1.04$, with a p-value less than $1.e^{-87}$. This empirical distribution can be approximated by a Laplace distribution (with $\lambda = 0.78$):

$$f_{TECU}(TECU = t) \propto e^{-1.27t} \tag{22}$$

The linear law fails for values of RMSE greater than 10 TECUs, which is explained by the fact that there are extremely few examples of errors greater than this value. The linear approximation also fails values of RMSE at lower end, that is < 0.75 TEC values. The empirical cumulative fraction of errors

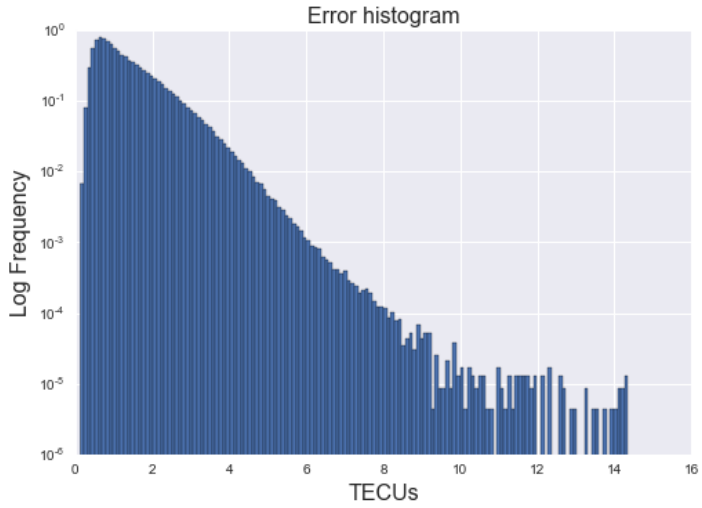

Figure 17. Normalized Log frequency histogram of the forecasting RMSE.

greater than a given TECU value are presented in Table 4. Interestingly, the accumulated probability diminishes very fast, and $Prob(\tau > 4TECUs)$ is lower than 1 in 65 pixels, and $Prob(\tau > 8TECUs)$ is lower than 1 in 8000 pixels. However, as explained in Section 5.3, the extreme values of the errors are isolated in the borders of the highly ionized regions.

6. Conclusions

In this paper, we have presented a forecasting method for TEC maps, based on local information that model possible local distortions. In addition to using several previous maps to make the prediction, we make a decomposition in the space spanned by the components in tangent manifold to each maps. The physical justification, is that the change at horizons less than 6 hours will consist of small changes along the trajectory that is followed by the sequence. The performance has been acceptable for forecast horizons up to 6 hours. For a horizon at 24 hours, the performance does not differ much from freezing the maps, which indicates both, that the hypothesis underlying the model of tangent spaces does not hold up to this time horizon, and that also that a simple linear combination of past observations is not enough for modeling a TEC map at a day's horizon. This is explained because at this horizon, the changes in ionization have external origins that cannot be derived from the time series.

Author Contributions: Methodology, E.M.M., A.G.R, M.H.P. and H.Y.; Software, E.M.M. and A.G.R; Validation, E.M.M., A.G.R, M.H.P. and H.Y.; Formal Analysis, E.M.M., A.G.R, M.H.P. and H.Y.; Data Curation, A.G.R; Writing—Review & Editing, E.M.M. and A.G.R.

Funding: This research was funded by the project TEC2015-69266-P (MINECO/FEDER, UE).

Acknowledgments: The authors are also grateful to the International GNSS Service (IGS; [36]) for providing GNSS open-data and ionospheric open-products.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. García Rigo, A., Monte Moreno,E., Hernández Pajares,M., Juan,J. M., Sanz,J., Aragón Angel,A. and Salazar,D. Global prediction of the vertical total electron content of the ionosphere based on GPS data. *Radio science* 46, no. 6. 2011
2. Béniguel, Yannick, Iurii Cherniak, Alberto Garcia-Rigo, Pierrick Hamel, Manuel Hernández-Pajares, Roland Kameni, Anton Kashcheyev et al. MONITOR Ionospheric Network: two case studies on scintillation and electron content variability. *Annales Geophysicae*, vol. 35, no. 3, p. 377. Copernicus GmbH, 2017.
3. Silvestrin, Pierluiggi, Michael Berger, Yann Kerr, and Jordi Font. ESAs second earth explorer opportunity mission: The soil moisture and ocean salinity mission-SMOS. *IEEE Geoscience and Remote Sensing Newsletter* 118 (2001): 11-14.
4. García-Rigo, A., D. Roma-Dollase, M. Hernández-Pajares, H. Lyu, Z. Li, N. Wang, M. Terkildsen et al. Contributions to real time and near real time Ionosphere Monitoring by IAGs RTIM-WG. *IAG-IASPEI Joint Scientific Assembly*. 2017.
5. Status of the IGS ionosphere products and future developments. *In IGS Analysis Center Workshop* (2008)
6. Schaer, S., and Société helvétique des sciences naturelles. Commission géodésique. Mapping and predicting the Earth's ionosphere using the Global Positioning System. *Institut fur Geodesie und Photogrammetrie, Eidg. Technische Hochschule Zürich*. (Vol. 59)
7. Cander, Lj R., M. M. Milosavljevic, S. S. Stankovic, and S. Tomasevic. Ionospheric forecasting technique by artificial neural network. *Electronics Letters* 34, no. 16 (1998): 1573-1574.
8. Francis, N. M., P. S. Cannon, A. G. Brown, and D. S. Broomhead. Nonlinear prediction of the ionospheric parameter foF2 on hourly, daily, and monthly timescales. *Journal of Geophysical Research: Space Physics* 105, no. A6 (2000): 12839-12849.
9. Tulunay, Ersin, Erdem Turker Senalp, Sandro Maria Radicella, and Yurdanur Tulunay. Forecasting total electron content maps by neural network technique. *Radio science* 41, no. 4 (2006).
10. Muhtarov, Plamen, and Ivan Kutiev. Autocorrelation method for temporal interpolation and short term prediction of ionospheric data. *Radio Science* 34, no. 2 (1999): 459-464.
11. Dick, M. I., M. F. Levy, L. R. Cander, I. Kutiev, and P. Muhtarov. Short-term ionospheric forecasting over Europe. *IEE National Conference on Antennas and Propagation*, 1999. pp. 105-107. IET, 1999.
12. Stanislawska, Iwona, and Zbigniew Zbyszynski. Forecasting of the ionospheric quiet and disturbed foF2 values at a single location. *Radio Science* 36, no. 5 (2001): 1065-1071.
13. Muhtarov, P., I. Kutiev, Lj R. Cander, B. Zolesi, G. de Franceschi, M. Levy, and M. Dick. European ionospheric forecast and mapping. *Physics and Chemistry of the Earth, Part C: Solar, Terrestrial and Planetary Science* 26, no. 5 (2001): 347-351.
14. Krankowski, Andrzej, Nadegda P. Sergeenko, Lubomir W. Baran, Irk I. Shagimuratov, and Marina V. Rogova. The short-term forecasting of the total electron content.*Artificial Satellites* 40 (2005): 137-146.
15. Hernández-Pajares, M., Juan, J. M., and Sanz, J. New approaches in global ionospheric determination using ground GPS data. *Journal of Atmospheric and Solar-Terrestrial Physics*, 61(16), 1237-1247.
16. Orús, R., Hernández-Pajares, M., Juan, J. M., and Sanz, J. Improvement of global ionospheric VTEC maps by using kriging interpolation technique. *Journal of Atmospheric and Solar-Terrestrial Physics*, 67(16), 1598-1609.
17. Orús, R., Hernández-Pajares, M., Juan, J. M., Sanz, J., and García-Fernández, M. Performance of different TEC models to provide GPS ionospheric corrections. *Journal of Atmospheric and Solar-Terrestrial Physics*, 64(18), 2055-2062

18. Hernández Pajares, M., Juan, J. M., Sanz, J., and Colombo, O. L. Application of ionospheric tomography to real time GPS carrier phase ambiguities resolution, at scales of 400 - 1000 km and with high geomagnetic activity. *Geophysical Research Letters*, 27(13), 2009-2012.
19. Roma-Dollase, David, Manuel Hernández-Pajares, Andrzej Krankowski, Kacper Kotulak, Reza Ghoddousi-Fard, Yunbin Yuan, Zishen Li et al. Consistency of seven different GNSS globl ionospheric mpping techniques during one solr cycle. *Journal of Geodesy* (2017): 1-16.
20. García Rigo, A. (2012). Contributions to ionospheric determination with global positioning system: solar flare detection and prediction of global maps of total electron content. *PhD Thesis, UPC*
21. Hernández−Pajares, M., J. M. Juan Zornoza, J. Sanz Subirana, R. Farnworth, and S. Soley.(2011) EGNOS Test Bed Ionospheric Corrections Under the October and November 2003 Storms, *IEEE Transactions On Geoscience And Remote Sensing*, VOL. 43,NO. 10, OCTOBER 2005."
22. Venkateswarlu, G., and Sarma, A. D. (2017). A New Technique Based on Grey Model for Forecasting of Ionospheric GPS Signal Delay Using GAGAN Data. *Progress In Electromagnetics Research, 59, 33-43*.
23. Huang, Z., Li, Q. B., and Yuan, H. (2015). Forecasting of ionospheric vertical TEC 1-h ahead using a genetic algorithm and neural network. *Advances in Space Research, 55(7), 1775-1783*.
24. Meng, X., Mannucci, A. J., Verkhoglyadova, O. P., and Tsurutani, B. T. (2016). On forecasting ionospheric total electron content responses to high-speed solar wind streams. *Journal of Space Weather and Space Climate, 6, A19.*
25. Feltens, J., M. Angling, N. Jackson-Booth, N. Jakowski, M. Hoque, M. Hernández Pajares, A. Aragón Angel, R. Orús, and R. Zandbergen. Comparative testing of four ionospheric models driven with GPS measurements. *Radio Science* 46, no. 6 (2011).
26. Hernández-Pajarez, Manuel, David Roma Dollase, Andrzej Krankowski, Alberto García Rigo, and Raul Orús Pérez Comparing performances of seven different global VTEC ionospheric models in the IGS context. *International GNSS Service Workshop (IGS 2016): Sydney, Australia* february 8-12, 2016, pp. 1-13. International GNSS Service (IGS), 2016.
27. Gonzalez, R. C., and Woods, R. E (2008). Digital image processing. *Addison-Welsley*
28. Simard, P., LeCun, Y. and Denker, J.S., (1993). Efficient pattern recognition using a new transformation distance. *Advances in neural information processing systems*, pp. 50-58.
29. Simard, P., LeCun, Y., Denker, J., and Victorri, B. (1998). Transformation invariance in pattern recognition tangent distance and tangent propagation. *Neural networks: tricks of the trade*, pp. 549-550.
30. Hastie, T. and Simard, P (1995). Learning prototype models for tangent distance. *Advances in neural information processing systems* , pp. 999-1006.
31. Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep learning (adaptive computation and machine learning series). *Adaptive Computation and Machine Learning series*, 800.
32. Friedman, J., Hastie, T. and Tibshirani, R. (2001). The elements of statistical learning. *Springer series in statistics.*, 800.
33. Mills, Terence C (1991). Time series techniques for economists. *Cambridge University Press*
34. (2018), Matplotlib documentation, Available from: https://matplotlib.org/api/_as_gen/matplotlib.pyplot.boxplot.html (Accessed 1 Februrary 2018)
35. Lyu, Haixia, Manuel Hernández-Pajares, Metin Nohutcu, Alberto Garcá-Rigo, Hongping Zhang, and Jingnan Liu. The Barcelona Ionospheric Mapping Function (BIMF) and its application to northern mid-latitudes. *GPS Solutions* 22 (2018): 1-13.
36. Dow, J. M., Neilan, R., and Rizos, C.: The international GNSS service in a changing landscape of global navigation satellite systems, *Journal of Geodesy*, 83, 191–198, 2009.